

**Scene level**HM3D household scenes  $\mathcal{S}$  $|\mathcal{S}| = 181$  scenes rumah tangga dengan tata letak, furnitur, dan konfigurasi objek yang beragam.**Task level**Generator tugas  $G_{\text{task}} : \mathcal{S} \rightarrow \mathcal{T}$  $\mathbb{E}[|\mathcal{T}_s|] \approx 5$  tugas per scene  $\Rightarrow |\mathcal{T}| \approx 905$  tugas LH-VLN.

2–6 subtugas (aksi simbolik: Move, Grab, Release).

Target sebaran subtugas per tugas: 70%: 3–4, 25%: 2, 5%: 5–6.

**Episode level (eksekusi di Habitat)**

Agen referensi mengeksekusi setiap tugas beberapa kali hingga batas percobaan tertentu.

Batas keras: 500 langkah simulasi per episode dan batas waktu *wall-clock*.Target:  $\approx 500$  episode sukses dengan panjang rata-rata 150–250 langkah, dan setidaknya sepertiga episode sukses memiliki trajektori  $> 200$  langkah untuk menjamin sifat *long-horizon*.**Himpunan trajektori sukses  $D_{\text{traj}}$** Episode yang berhasil menyelesaikan semua subtugas dalam batas langkah dan waktu dimasukkan ke  $D_{\text{traj}}$ .

Episode gagal hanya digunakan untuk statistik kegagalan (mis. tingkat kesuksesan per tugas), tidak masuk ke tahap segmentasi.

**Segmentation level**Modul segmentasi  $S_{\text{seg}} : D_{\text{traj}} \rightarrow D^*$ 

Trajektori sukses dipecah menjadi beberapa segmen berdasarkan perubahan target lokal dan pola aksi (navigasi dan manipulasi).

Dengan rata-rata 3–4 subtugas per tugas, satu episode sukses menghasilkan  $\approx 4$ –6 segmen.**Final dataset  $D^*$** Dari  $\approx 500$  episode sukses dan  $\approx 4$ –6 segmen per episode, diperoleh sekitar 2,000–3,000 segmen.Setiap segmen dilengkapi instruksi navigasi bilingual, sehingga  $D^*$  cukup besar untuk (a) analisis statistik *code-switching* di tingkat korpus, dan (b) evaluasi variasi pola bahasa lintas scene, konfigurasi robot, dan kompleksitas tugas.